



SADA

South African Data Archive

**GENERAL HOUSEHOLD SURVEY, 2003
(South Africa)**

Statistics South Africa (Stats SA)

CODEBOOK

SADA 0136

BIBLIOGRAPHIC CITATION

Any publication or other presentation based in whole or part on the data and documentation supplied by SADA must prominently use the following citation:

Statistics South Africa. *General Household Survey, 2003* [Computer file]. S0136. Pretoria: Statistics South Africa [producer], 2003. Pretoria: South African Data Archive, National Research Foundation [distributor], 2004.

DEPOSIT OF PUBLICATIONS

At least one copy of any published work or report based in whole or part on the dataset will be deposited with the South African Data Archive. Please indicate the title and number of the study utilised.

LIMITATION OF LIABILITY

The Archive and the depositor of the dataset supplied to the user bear no legal responsibility for their accuracy or comprehensiveness.

| | |
|------------------|---|
| Postal Address: | SOUTH AFRICAN DATA ARCHIVE NATIONAL RESEARCH FOUNDATION P.O. BOX 2600 Pretoria 0001 |
| Telephone: | +27 12 481 4158/ 4120 |
| Telefax: | +27 12 349 1179 |
| Electronic mail: | sada@nrf.ac.za |
| World Wide Web: | http://www.nrf.ac.za/sada |

STUDY DESCRIPTION

SADA 0136

TITLE: General Household Survey, 2003 (GHS 2).

PRINCIPAL INVESTIGATOR: Statistics South Africa (Stats SA)

DEPOSITOR: Statistics South Africa

ORGANISATION HOUSING THE DATA:

Statistics South Africa
Private Bag x44
Pretoria, 0001
South Africa
Tel: (012) 310 8600
Fax: (012) 310 8500
Email: info@statssa.gov.za
Web: <http://www.statssa.gov.za>

KEYWORDS:

Employment; Unemployment; Household surveys; Housing; Education and Health.

SERIES DATA:

| | | |
|--------------------------|-------------------------|----------------------|
| Series name: | Principal Investigator: | Currently available: |
| General Household Survey | Statistics South Africa | 2002 - 2011 |

PURPOSE:

The GHS is an annual household survey, specifically designed to measure various aspects of the living circumstances of South African households.

GEOGRAPHIC LOCATION: South Africa

IMPORTANT VARIABLES: Employment, unemployment, education, health, access to facilities.

DEMOGRAPHIC VARIABLES: Age, gender, level of education, marital status, use of health services, economic activity and employment.

UNIVERSE: Households in the nine provinces of South Africa

METHOD OF DATA COLLECTION: Survey Questionnaire

SAMPLING:

A multi-stage stratified sample was drawn using probability proportional to size principles.

The sample was drawn from the master sample, which Statistics South Africa uses to draw samples for its regular household surveys. The master sample is drawn from the database of enumeration areas (EAs) established during the demarcation phase of Census 1996. As part of the master sample, small EAs consisting of fewer than 100 households are combined with adjacent EAs to form primary sampling units (PSUs) of at least 100 households, to allow for repeated sampling of dwelling units within each PSU. The sampling procedure for the master sample involves explicit stratification by province and within each province, by urban and non-urban areas. Within each stratum, the sample was allocated disproportionately. A PPS sample of PSUs was drawn in each stratum, with the measure of size being the number of households in the PSU. Altogether approximately 3 000 PSUs were selected. In each selected PSU a systematic sample of ten dwelling units was drawn, thus, resulting in approximately 30 000 dwelling units. All households in the sampled dwelling units were enumerated.

The master sample is divided into five independent clusters. In order to avoid respondent fatigue (the LFS is a rotating panel survey which is conducted twice yearly), the GHS sample uses a different cluster from the LFS clusters.

WEIGHTING:

A two-stage weighting procedure was done on the GHS 2003 that resulted in two sets of weights, household weights and person weights. The household weights are used when analysis is at the household level and the person weights are used when analysis is at the individual level.

UNITS OF OBSERVATION: Households

DATE OF DATA COLLECTION: July 2003

EXTENT OF DATA COLLECTION: 5 data files in ASCII and machine-readable documentation.

FILE SPECIFICATIONS:

Part 1: House Data

| | |
|-----------------------------|----------|
| Number of cases: | 26 398 |
| Number of records: | 26 398 |
| Number of records per case: | 1 |
| Logical Record Length: | 230 |
| Number of variables: | 156 |
| Number of kilobytes: | 5 956 KB |

Part 2: Person Data

| | |
|-----------------------------|-----------|
| Number of cases: | 99 428 |
| Number of records: | 99 428 |
| Number of records per case: | 1 |
| Logical Record Length: | 134 |
| Number of variables: | 92 |
| Number of kilobytes: | 13 109 KB |

Part 3: stratpsu data

| | |
|-----------------------------|-------|
| Number of cases: | 2 961 |
| Number of records: | 2 961 |
| Number of records per case: | 1 |
| Logical Record Length: | 18 |
| Number of variables: | 3 |
| Number of kilobytes: | 58 KB |

Part 4: Tourism Data

| | |
|-----------------------------|----------|
| Number of cases: | 26 397 |
| Number of records: | 26 397 |
| Number of records per case: | 1 |
| Logical Record Length: | 182 |
| Number of variables: | 71 |
| Number of kilobytes: | 4 718 KB |

Part 5: Worker Data

| | |
|-----------------------------|----------|
| Number of cases: | 68 245 |
| Number of records: | 68 245 |
| Number of records per case: | 1 |
| Logical Record Length: | 93 |
| Number of variables: | 46 |
| Number of kilobytes: | 6 265 KB |

PUBLICATIONS:

Statistics South Africa. 2003. General Household Survey, 2003. Pretoria: Statistics South Africa.

NOTES:

Most questions in the General Household Survey questionnaire are pre-coded, i.e. there is a set number of choices from which one or more must be selected. For open-ended 'write-in' questions, the description will note that post-coding occurred and explain how this was done. For most variables the coding is apparent from the questionnaire (available elsewhere in the documentation) and is not repeated in the variable description. Where the coding is not apparent, the description either provides the codes or indicates where code lists are to be found.

RESTRICTIONS:

Only available to bona fide researchers.

DATE STUDY ADDED:

01 November 2004